(12)

STATISTICAL EFFICIENCY OF REGENERATIVE SIMULATION METHODS FOR NETWORKS OF QUEUES

by

Donald L. Iglehart and Gerald S. Shedler

TECHNICAL REPORT NO. 58

May 1982

Prepared under Contract N00014-76-C-0578 (NR 042-343)*

for the

Office of Naval Research

DEPARTMENT OF OPERATIONS RESEARCH
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

*This research was also partially supported under National Science Foundation Grant MCS79-09139.

1. INTRODUCTION

When developing simulation methodology it is important to be able to assess the statistical efficiency of proposed estimation procedures. This paper is concerned with the assessment of the statistical efficiency of regenerative methods for simulation of networks of queues with general service times. Our estimation procedures are applicable to networks that have a "single state" for passage times; regenerative cycles are defined in terms of the single state. We provide a comparison of two methods for estimation of passage time characteristics in closed networks with priorities among job classes. Passage times (informally the times for a job to traverse a portion of a network) are important in computer and communication system models where they represent job "response times."

The marked job method (Iglehart and Shedler [9]) prescribes observation of passage times for an arbitrarily chosen, distinguished job. With the broadly applicable marked job method the half-length of the confidence interval (obtained from a simulation of fixed length) for the expected value of a general function f of the limiting passage time is proportional to a certain quantity $e(f)$. The labelled jobs method (Shedler and Southard [12]) provides estimates for passage times through a subnetwork. With the labelled jobs method, observed passage times for all the jobs are used to construct point and interval estimates and (with the same constant of proportionality) the half-length of the confidence interval is proportional to a quantity $e^0(f)$. Since these quantities are independent of the blocks of the underlying regenerative process, they are appropriate measures of the statistical efficiency of the estimation procedures. For Markovian networks of queues, using computational results of Hordijk, Iglehart, and Schassberger [5], it is possible to compute theoretical values for expected passage times and the associated variance constants appearing in central limit theorems (c.l.t.'s) used to form confidence

intervals; here f is the identity function. This leads to a quantitative assessment (Iglehart and Shedler [8]) of the relative statistical efficiencies of the estimation procedures (for expected passage times) in [7] for networks with Cox-phase service times.

For networks of queues with general service times, there is little hope of computing the needed theoretical values, even for expected passage times. Using central limit theorem and continuous mapping theorem arguments, in this paper we show that for any function f (and all numbers of jobs in the network) $e^0(f) \le e(f)$; i.e., the confidence intervals constructed using the labelled jobs method are shorter than those obtained from the marked job method. This is consistent with intuition since the labelled jobs method extracts more passage time information from a fixed length simulation run.

2. CLOSED, MULTICLASS NETWORKS OF QUEUES AND PASSAGE TIMES

We consider closed networks of queues having a finite number of *jobs* (customers), N, a finite number of (single or multiple server) *service centers*, s, and a finite number of (mutually exclusive) *job classes*, c. At every epoch of continuous time each job is in exactly one job class, but jobs may change class as they traverse the network. Upon completion of service at center i a job of class j goes to center k and changes to class l with probability $p_{ij,kl}$, where

$$P = \left\{ p_{ij,kl}\colon (i,j),(k,l) \in C \right\}$$

is a given irreducible Markov matrix and $C \subseteq \{1,2,...,s\} \times \{1,2,...,c\}$ is the set of (center, class) pairs in the network. Note that in accordance with the matrix P some centers may never see jobs of certain classes. At each service center jobs queue and receive service according to a fixed priority scheme among classes; the priority scheme may differ from center to center. Within a class at a center, jobs receive service according

to a fixed queue service discipline, e.g., first-come, first-served (FCFS). According to a fixed procedure for each center, a job in service may or may not be preempted if another job of higher priority joins the queue at the center. (For technical reasons, we assume that any interruption of service is of the preemptive-repeat type.) A job that has been preempted receives additional service at the center before any other job of its class at the center receives service.

All service times are assumed to be mutually indeoendent. We also suppose that service times at any center have finite mean but otherwise arbitrary density function which is continuous and positive on $(0,\infty)$. Parameters of the service time distribution may depend on the service center, the class of job in service, and the "state" (as defined in Equation (2.1) below) of the entire network at the time service begins. In order to characterize the state of the network at time t, we let $S_i(t)$ denote the class of the job in service at center i at time t, where $i = 1,2,...,s$; by convention $S_i(t) = 0$ if at time t there are no jobs at center i. If center i has more than one server, we enumerate the servers at center i and let $S_i(t)$ be a vector which records the class of the job receiving service from each server at the center. (A job receives service from the lowest numbered available server.) Denoting the number of job classes serviced at center i by $k(i)$ the classes of jobs serviced at center i ordered by decreasing priority are $j_1(i),j_2(i),...,j_{k(i)}(i)$, elements of the set $\{1,2,...,c\}$. We denote by $C_{j_1}^{(i)}(t),...,C_{j_{k(i)}}^{(i)}(t)$ the number of jobs in queue at time t of the various classes of jobs serviced at center i, $i = 1,2,...,s$.

According to the following scheme, we order the N jobs in a linear stack. For $t \geq 0$ we set

$$Z(t) = (C_{j_{k(1)}}^{(1)}(t),...,C_{j_1}^{(1)}(t),S_1(t);...;C_{j_{k(s)}}^{(s)}(t),...,C_{j_1}^{(s)}(t),S_s(t)). \tag{2.1}$$

The *job stack at time t* then corresponds to the order of components in the vector $Z(t)$ after ignoring any zero components. Within a class at a particular service center, jobs waiting appear in the job stack in FCFS order; i.e., in order of their arrival at the center, the latest to arrive being closest to the top of the stack. The *job stack process* $Z = \{Z(t):t\geq 0\}$ has a finite state space, D^*.

For $t\geq 0$ denote by $N(t)$ the position (from the top) of the marked job in the job stack at time t. Then set

$$X(t) = (Z(t), N(t)) \tag{2.2}$$

and call $X = \{X(t):t\geq 0\}$ the *augmented job stack process.* Passage times are specified in terms of the marked job by means of four subsets (A_1, A_2, B_1, and B_2) of the state space, E^*, of the stochastic process X. The sets A_1, A_2 [resp. B_1, B_2] jointly define the random times at which passage times for the marked job start [resp. terminate]. The sets A_1, A_2, B_1, and B_2 in effect determine when to start and stop the clock measuring a particular passage time of the marked job. We assume that the start and termination times for the specified passage time strictly alternate.

In terms of the sets A_1, A_2, B_1, and B_2, we define two sequences of random times, $\{S_j:j\geq 0\}$ and $\{T_j:j\geq 1\}$, where S_{j-1} is the start time of the jth passage time for the marked job and T_j is the termination time of this jth passage time. Assuming that the initial state of the process X is such that a passage time for the marked job begins at $t=0$, let

$$S_0 = 0$$

$$S_j = \inf\{\tau_n \geq T_j: X(\tau_n) \in A_2, X(\tau_{n-1}) \in A_1\}, j \geq 1$$

and

$$T_j = \inf\{\tau_n \geq S_{j-1}: X(\tau_n) \in B_2, X(\tau_{n-1}) \in B_1\}, j \geq 1,$$

where $\{\tau_n : n \geq 0\}$ are the jump times of the process X. Then the jth passage time for the marked job is $P_j = T_j - S_{j-1}$, $j \geq 1$.

Let $U(z)$ be the set of all $(i,j) \in C$ such that in state $z \in D^*$ there is a job of class j in service at center i. For $z, z' \in D^*$ and $u = (i,j) \in U(z)$, let $q(z';z,u)$ be the probability that the job stack process Z jumps (in one step) to state z', given that in state z there is a completion of service to a job of class j at center i. We write $z \to z'$ when $q(z';z,u) > 0$ for some $u \in U(z)$. For all $z, z' \in D^*$, we say that z' *is accessible from* z and write $z \rightsquigarrow z'$ if there exists a finite sequence $u_0', z_1, u_1', \ldots, u_n'$ (center, class) pairs and job stacks such that

$$q(z_1;z,u_0')q(z_2;z_1;u_1')\ldots q(z';z_n,u_n') > 0. \tag{2.3}$$

When $z \rightsquigarrow z'$ and $z' \rightsquigarrow z$ we say that z *and* z' *communicate* and write $z \leftrightsquigarrow z'$.

We also define $U(x)$ for $x \in E^*$: $U(x) = U(z)$ when $x = (z,n)$ for some $z \in D^*$ and $n \in \{1,2,\ldots,N\}$. For $x, x' \in E^*$ and $u = (i,j) \in U(x)$, we denote by $p(x';x,u)$ the probability that the augmented job stack process X jumps to state x', given that in state x there is a completion of service to a job of class j at center i. We write $x \to x'$ when $p(x';x,u) > 0$ for some $u \in U(x)$. We say that x' *is accessible from* x and write $x \rightsquigarrow x'$ if there exists a finite sequence $u_0', x_1, u_1', \ldots, u_n'$ of (center, class) pairs and augmented job stacks such that

$$p(x_1;x,u_0')p(x_2;x_1,u_1')\ldots p(x';x_n,u_n') > 0. \tag{2.4}$$

When $x \rightsquigarrow x'$ and $x' \rightsquigarrow x$ we say that x *and* x' *communicate* and write $x \leftrightsquigarrow x'$.

3. SIMULATION FOR PASSAGE TIMES

In the absence of some restriction on the building blocks of a network of queues with priorities among job classes, the sequence $\{P_j : j \geq 1\}$ of passage times for the marked job need not converge in distribution to a random variable independent of the initial state of the system. We make the further assumption that for some $z^* \in D^*$ the sets $D = \{z \in D^* : z^* \rightsquigarrow z\}$ and $E = \{(z,n) \in E^* : z \in D\}$ are irreducible; i.e., $z \leftrightsquigarrow z'$ for all $z, z' \in D$ and $x \leftrightsquigarrow x'$ for all $x, x' \in E$.

For networks with two or more service centers ($s>1$), the sets D and E are irreducible (cf. Shedler and Slutz [10]) if for some service center i_0 either $k(i_0) = 1$ (only one class served by center i_0) or service at center i_0 to a job of class j_{i_0} (the lowest priority job class seen by center i_0) is preempted when a job of higher priority joins the queue. Let $z^* = z^*_{i_0}$, the job stack in which there is one job of class $j_{k(i_0)}(i_0)$ in service at center i_0 and N-1 jobs of class $j_{k(i_0)}(i_0)$ in queue at center i_0 (or in service if center i_0 is a multiple server center). Then the sets $D = \{z \in D^* : z^*_{i_0} \rightsquigarrow z\}$ and $E = \{(z,n) \in E^* : z \in D\}$ are irreducible.

Label the jobs from 1 to N and for $t \geq 0$ denote by $N^i(t)$ the label of the job in position i of the job stack at time t, $1 \leq i \leq N$. Set

$$X^0(t) = (Z(t), N^1(t), \ldots, N^N(t))$$

and call X^0 the *fully augmented job stack process*. It can be shown (cf. Appendix 1 of [7]) that $\{P^0_n : n \geq 1\}$, the sequence of passage times (irrespective of job identity) enumerated in termination order, converges in distribution to a random variable P^0. Moreover $P^0 = P$, the limiting passage time for any marked job. The goal of the simulation is estimation of

$$r(f) = E\{f(P)\},$$

where f is a real-valued (measurable) function. We assume that $E\{|f(P)|\}<\infty$ and $P\{P\in D(f) = 0\}$, where $D(f)$ is the set of discontinuities of the function f.

The marked job method

Define a set S according to

$$S = \{(k,m): k\in A_1,\ m\in A_2 \text{ and } p(m;k,u)>0 \text{ for some } u\in U(k)\}. \tag{3.1}$$

For $(k,m)\in S$ the entrances of X to state m from state k correspond to the starts of passage times for the marked job. For $z\in D$ and $n\in\{1,2,\ldots,N\}$, we write $h(z,n) = (i,j)$ when the job in position n in the job stack associated with state z is of class j at center i. Now define a subset S' of S according to

$$S' = \{(z,N,z',n')\in S: \text{ for some single server center } i \text{ and some } (i,j_l(i))\in C,$$
$$h(z,N) = (i,j_l(i)) \text{ and } h(z,n) = (i,j_{l_n}(i)) \text{ with } l_n\ge l,\ 1\le n<N\} \tag{3.2}$$

and assume that $S'\neq\phi$. This condition ensures that there exists a *single state* $((z,N))$ *of the augmented job stack process* X such that the marked job is in service at some single server center, the other N-1 jobs are in queue at the center as jobs of equal or lower priority, and with positive probability a passage time for the marked job starts upon completion of the service in progress. For $(z,N,z',n')\in S'$ the entrances of X to state (z',n') from single state (z,N) correspond to starts of passage times for the marked job. Also define

$$A_2' = \{(z',n')\in A_2: (z,N,z',n')\in S' \text{ for some } z\in D\}. \tag{3.3}$$

Then based on a representation of X as an irreducible generalized semi-Markov process in the sense of Whitt [13], it can be shown ([9], Proposition (3.1)) that $P\{X(S_n) = x' \text{ i.o.}\} = 1$ for any $x' \in A_2'$.

Select $x' \in A_2'$, begin the simulation of X with $X(0) = x'$, and carry out the simulation of X in cycles defined by the successive entrances of $\{X(S_n):n\geq 0\}$ to x'. Let α_m denote the length (in discrete time units) of the mth cycle of $\{X(S_n):n\geq 0\}$ and define $\beta_0 = 0$ and $\beta_m = \alpha_1 + \dots + \alpha_m$, $m\geq 1$; α_m is the number of passage times for the marked job in the mth cycle. Also define

$$Y_m(f) = \sum_{j=\beta_{m-1}+1}^{\beta_m} f(P_j)$$

for any function f.

By Proposition (3.4) of [9], the process $\{(X(S_n),P_{n+1}):n\geq 0\}$ is a regenerative process in discrete time. It follows that the sequence of pairs of random variables $\{(Y_m(f),\alpha_m):m\geq 1\}$ are independent and identically distributed, and since $E\{\,|f(P)|\,\}<\infty$ by assumption,

$$E\{f(P)\} = E\{Y_1(f)\}/E\{\alpha_1\}.$$

With these results, the standard regenerative method (Crane and Iglehart [4]) applies and (from a fixed number, n, of cycles) provides the strongly consistent point estimate $\hat{r}_n(f) = \overline{Y}_n(f)/\overline{\alpha}_n$ for $r(f)$, where

$$\overline{Y}_n(f) = n^{-1}\sum_{m=1}^{n} Y_m(f)$$

and

$$\bar{\alpha}_n = n^{-1} \sum_{m=1}^{n} \alpha_m.$$

Confidence intervals for $r(f)$ are based on the c.l.t.

$$n^{1/2} \left[\hat{r}_n(f) - r(f)\right] / \left[(\sigma(f)/E\{\alpha_1\}\right] \Rightarrow N(0,1) \tag{3.4}$$

as $n \to \infty$, where $\sigma^2(f)$ (assumed finite) is the variance of $Y_1(f) - r(f)\alpha_1$ and $N(0,1)$ is a standardized (mean 0, variance 1) normal random variable.

The labelled jobs method

Analogous to the set

$$S = \{(k,m):\ k \in A_1,\ m \in A_2 \text{ and } p(m;k,u) > 0 \text{ for some } u \in U(k)\}$$

of Equation (3.1), define a set T according to

$$T = \{(k,m):\ k \in B_1,\ m \in B_2 \text{ and } p(m;k,u) > 0 \text{ for some } u \in U(k)\}. \tag{3.5}$$

For $(k,m) \in T$ the entrances of X to state m from state k correspond to the terminations of passage times for the marked job. Let $H = H_1 \cup H_2 \subseteq C$, where (i,j) is in H_1 [resp. H_2] if it is possible for the marked job to be of class j at center i when the passage time specified by the sets A_1, A_2, B_1, and B_2 terminates [resp. is not underway]. The labelled jobs method applies to passage times for which $S \cap T = \phi$ so that the set H is nonempty. (Passage times for which $S \cap T = \phi$ correspond to passage of a job through a subnetwork.)

An element z of the set D is called a *single state of the job stack process* for the passage time specified by the sets A_1, A_2, B_1, and B_2 if

(i) $h(z,n) = (i_0, j_{k_n}(i_0)) \in H$ for some single server center i_0, $n = 1,2,\ldots,N$;

and

(ii) there exists $(z_1,m) \in B_1$ such that $(z_1,m) \to (z,n)$ for some $(z,n) \in B_2$.

According to this definition a single state of the job stack process is a configuration, z, of the job stack such that no passage times are underway ($h(z,n) \in H$ for all n), all jobs are at the same center (i_0) with exactly one job in service, and there exis such that a passage time for some job terminates when the job stack process jumps to z. We assume that a single state of the job stack process exists.

Select a single state, z_0, of the job stack process and an initial state $(z_0, n^1, \ldots, n^N)$ for the fully augmented job stack process X^0. Let T_n^0 be the termination time of P_n^0, $n \geq 1$. Denote by $\{\beta_k^0 : k \geq 1\}$ the indices of the successive passage times (irrespective of job identity) which terminate with the job stack process in state z_0. Let $T_0^0 = \beta_0^0 = 0$. Carry out the simulation of the process X^0 in blocks defined by the successive epochs $\{T_{\beta_k^0}^0 : k \geq 1\}$. Set

$$Y_m^0(f) = \sum_{j=\beta_{m-1}^0+1}^{\beta_m^0} f(P_j^0)$$

and $\alpha_m^0 = \beta_m^0 - \beta_{m-1}^0$, $m \geq 1$.

It can be shown ([12], Proposition (3.2) and (3.4)) that $P\{Z(T_n^0) = z \text{ i.o.}\} = 1$ for any single state z of the job stack process, and that the process $\{(Z(T_n^0), P_{n+1}^0) : n \geq 0\}$ is a regenerative process in discrete time. It follows that the pairs of random variables $\{(Y_m^0(f), \alpha_m^0) : m \geq 1\}$ are independent and identically distributed, and since $E\{|f(P^0)|\} < \infty$ by assumption,

$$E\{f(P^0)\} = E\{Y_1^0(f)\}/E\{\alpha_1^0\}.$$

Then the standard regenerative method applies and (from a fixed number, n, of blocks) provides the strongly consistent point estimate

$$\hat{r}_n^0(f) = \overline{Y}_n^0(f)/\overline{\alpha}_n^0 = \sum_{m=1}^{n} Y_m^0(f) \Big/ \sum_{m=1}^{n} \alpha_m^0$$

for $r(f)$. Confidence intervals for $r(f)$ are based on the c.l.t.

$$n^{1/2}\left[\hat{r}_n^0(f) - r(f)\right] \Big/ \left[\sigma^0(f)/E\{\alpha_1^0\}\right] \to N(0,1), \tag{3.6}$$

where $(\sigma^0(f))^2$ (assumed finite) is the variance of $Y_1^0(f)-r(f)\alpha_1^0$.

(3.7) EXAMPLE. Consider a network with two service centers and one job class. Let $C = \{(1,1),(2,1)\}$ and suppose that the irreducible routing matrix P is

$$P = \begin{bmatrix} p & 1-p \\ 1 & 0 \end{bmatrix}.$$

(Thus with probability p a job completing service at center 1 joins the tail of the queue at center 1 and with probability $1-p$ joins the queue at center 2. A job completing service at center 2 joins the tail of the queue at center 1.) Taking into account the fixed number of jobs in the network, for $t \geq 0$ let $Z(t)$ be the number of jobs waiting or in service at center 1 at time t. Then for the job stack process $D^* = D = \{0,1,\ldots,N\}$ and for the augmented job stack process

$$E^* = E = \{(i,N)\colon 0 \leq i \leq N,\ 1 \leq n \leq N\}.$$

Denote by P the (limiting) passage time which starts when a job enters the center 1 queue upon completion of service at center 2 and terminates when the job joins the

center 2 queue. This passage time is specified by the sets $A_1 = \{(i,N):0\le i<N\}$, $A_2 = \{(i,1):0<i\le N\}$, $B_1 = \{(i,i):0<i\le N\}$, and $B_2 = \{(i,i+1):0\le i<N\}$. Both estimation methods are applicable. The set S correponding to the starts of passage times for the marked job is

$$S = \{(i,N,i+1,1):\ 0\le i<N\}$$

and the subset $S' = \{(0,N,1,1)\}$. There is one single state $((0,N))$ of the augmented job stack process and the set $A_2' = \{(1,1)\}$. For the labelled jobs method, the state $z_0=0$ is a single state of the job stack process.

4. COMPARISON OF METHODS

Assume that for the passage time P, both the marked job and labelled jobs methods are applicable; i.e., assume that the set S' of Equation (3.2) is nonempty, that the sets S and T of Equations (3.1) and (3.5) are disjoint, and that there exists a single state of the job stack process. For $t\ge 0$ let $L(t)$ be the last state occupied by the process X before jumping to $X(t)$. Similarly, let $L^0(t)$ be the last state occupied by the process X^0 be. jumping to $X^0(t)$. Set $V = \{V(t):t\ge 0\}$ and $V^0 = \{V^0(t):t\ge 0\}$, where $V(t) = (L(t),X(t))$ and $V^0(t) = (L^0(t),X^0(t))$. Denote the state space of V by F and the state space of V^0 by F^0. We work with the process V^0 and take job 1 to be the marked job. Selecting z_0, a single state of the job stack process, $v_0 = (z_0',m_1',z_0,m_1) \in T$, and $v_0^0 = (z_0',m_1',m_2',\ldots,m_N',z_0,m_1,m_2,\ldots,m_N) \in F^0$, we set $V^0(0) = v_0^0$.

Lemma (4.1) is the basis for construction of central limit theorems in continuous time for sequences of passage times. Let $\{\gamma_k^*:k\ge 0\}$ be a renewal process defined on the same sample space as $\{P_n^0:n\ge 1\}$ and set $\delta_{k+1}^* = \gamma_{k+1}^*-\gamma_k^*$. Let $\{P_{n_j}^0:j\ge 1\}$ be a subsequence

of $\{P_n^0:n\geq 1\}$ and denote by $m^*(t)$ the number of $P_{n_j}^0$ completed in $(0,t]$. Set $\alpha_k^* = m^*(\gamma_k^*)-m^*(\gamma_{k-1}^*)$ and

$$Y_k^*(f) = \sum_{j=m^*(\gamma_{k-1}^*)+1}^{m^*(\gamma_k^*)} f(P_{n_j}^0).$$

(4.1) LEMMA. Suppose that the pairs of random variables $\{(Y_k^*(f),\alpha_k^*):k\geq 1\}$ are independent and identically distributed and $r(f) = E\{Y_1^*(f)\}/E\{\alpha_1^*\}$. Then provided that $E\{\delta_1^*\}<\infty$, $E\{(\alpha_1^*)^2\}<\infty$, and $E\{(Y_1^*(|f|))^2\}<\infty$,

$$t^{1/2}\Big(\frac{1}{m^*(t)}\sum_{j=1}^{m^*(t)} f(P_{n_j}^0) - r(f)\Big)\Big/\Big[(E\{\delta_1^*\})^{1/2}\sigma^*(f)/E\{\alpha_1^*\}\Big] \to N(0,1)$$

as $t\to\infty$, where $(\sigma^*(f))^2 = \sigma^2(Y_1^*(f)-r(f)\alpha_1^*)<\infty$.

Proof: Let

$$A(t) = \sum_{j=1}^{m^*(t)} f(P_{n_j}^0) - r(f)m^*(t).$$

The Doeblin decomposition is given by

$$A(t) = \sum_{k=1}^{n^*(t)}\big[Y_k^*(f) - r(f)\alpha_k^*\big] + \sum_{j=m^*(\gamma_{n(t)}^*)+1}^{m^*(t)} f(P_{n_j}^0) - r(f)\big[m^*(t) - m^*(\gamma_{n^*(t)})\big], \quad (4.2)$$

where $\{n^*(t):t\geq 0\}$ is the counting process associated with the renewal process $\{\gamma_k^*:k\geq 0\}$. Now let $Z_k^*(f) = Y_k^*(f)-r(f)\alpha_k^*$ and observe that $E\{Z_k^*(f)\} = 0$ and the $Z_k^*(f)$'s are i.i.d. by assumption. The weak law for renewal counting processes yields $n^*(t)/t\to 1/E\{\delta_1^*\}$ as $t\to\infty$. Using these facts, the c.l.t. for random partial sums (cf. Chung [3], Theorem 7.3.2 and Billingsley [1], p. 16) and the continuous mapping theorem (cf. Billingsley [2], Theorem 25.7, Corollary 1) we obtain the c.l.t.

$$\sum_{k=1}^{n^*(t)} Z_k^*(f) \Big/ \left[\sigma^*(f)\, t^{1/2}/E\{\delta_1^*\}\right] \Rightarrow N(0,1) \tag{4.3}$$

as $t \to \infty$. From Equation (4.3) we see that, suitably normalized, the first term on the right-hand side of Equation (4.2) converges weakly to $N(0,1)$. The last two terms on the right-hand side of Equation (4.2) are remainder terms which, when divided by $t^{1/2}$, converge weakly to 0. This argument is standard (cf. Iglehart [6], p. 279) but requires $E\{(\alpha_1^*)^2\}$ and $E\{(Y_1^*(|f|))^2\}$ to be finite. Using the converging together theorem (cf. Billingsley [2], Theorem 25.4) and Equation (4.3) we have

$$A(t) \Big/ \left[\sigma^*(f)\, t^{1/2}/E\{\delta_1^*\}\right] \Rightarrow N(0,1) \tag{4.4}$$

as $t \to \infty$. Finally, dividing the numerator and denominator of the left-hand side of Equation (4.4) by $m^*(t)$, using the fact that $m^*(t)/t \to E\{\alpha_1^*\}/E\{\delta_1^*\}$ as $t \to \infty$ and the continuous mapping theorem we obtain the desired result. □

The labelled jobs method prescribes selection of a single state, z_0, of the job stack process along with $v_0 = (z_0', m_1', z_0, m_1) \in T$ and simulation of V^0 in blocks defined by exits from the set

$$\{(z_0', n_1', n_2', \ldots, n_N', z_0, n_1, n_2, \ldots, n_N) \in F^0 \colon (z_0', n_i', z_0, n_i) = v_0 \text{ for some } i,\ 1 \le i \le N\}.$$

By Lemma (4.1)

$$t^{1/2} \left(\frac{1}{m^0(t)} \sum_{n=1}^{m^0(t)} f(P_n^0) - r(f) \right) \Big/ \left[(E\{\delta_1^0\})^{1/2} \sigma^0(f) / E\{\alpha_1^0\} \right] \Rightarrow N(0,1)$$

as $t \to \infty$, where δ_1^0 is the length of a block and $m^0(t)$ is the number of passage times (irrespective of job identity) completed in the interval $(0,t]$. This c.l.t. implies that the

half-length of the confidence interval obtained from a simulation of fixed length is proportional to the quantity $e^0(f)$ defined by

$$e^0(f) = (E\{\delta_1^0\})^{1/2}\sigma^0(f)/E\{\alpha_1^0\}, \tag{4.5}$$

where $(\sigma^0(f))^2 = \sigma^2(Y_1^0(f)-r(f)\alpha_1^0)$ and $Y_1^0(f)$ are as in Section 3. (We assume that $E\{(\alpha_1^0)^2\}<\infty$ and $E\{(Y_1^0(|f|))^2\}<\infty$. This makes it possible to apply Lemma (4.1).) Since the numerator in this c.l.t. and the limit ($N(0,1)$) is independent of the state v_0 selected from T, so is the denominator; this is a consequence of the "convergence of types" theorem (cf. Billingsley [2], Theorem 14.2). Thus $e^0(f)$ is an appropriate measure of the statistical efficiency of the labelled jobs method.

The marked job method prescribes selection of $x' = (z',n') \in A_2'$ and simulation of V in blocks defined by the successive entrances to the set

$$\{(z_0,N,z^*,n^*) \in S' : z^* = z', n^* = n'\}.$$

By Lemma (4.1)

$$t^{1/2}\left(\frac{1}{m(t)}\sum_{n=1}^{m(t)} f(P_n) - r(f)\right)\Big/\left[(E\{\delta_1\})^{1/2}\sigma(f)/E\{\alpha_1\}\right] \Rightarrow N(0,1) \tag{4.6}$$

as $t\to\infty$, where δ_1 is the length of a block and $m(t)$ is the number of passage times completed in the interval $(0,t]$. By an argument similar to that used for the labelled jobs method, an appropriate measure of the statistical efficiency of the marked job method is the quantity

$$e(f) = (E\{\delta_1\})^{1/2}\sigma(f)/E\{\alpha_1\}, \tag{4.7}$$

where $(\sigma(f))^2 = \sigma^2(Y_1(f)-r(f)\alpha_1)$ and $Y_1(f)$ are as in Section 3. We shall show that $e^0(f) \le e(f)$ for all functions f.

In terms of the single state z_0 of the job stack process and the state $v_0 = (z_0', m_1', z_0, m_1) \in T$, define subsets $U^1, U^2, \ldots, U^N$ of F^0 by

$$U^i = \{(z_0', m_1', \ldots, m_i', n_{i+1}', \ldots, n_N', z_0, m_1, \ldots, m_i, n_{i+1}, \ldots, n_N) \in F^0\},$$

$1 \le i \le N$. According to this definition, entrance of V^0 to the set U^i corresponds to termination of a passage time for job 1 (in position m_1 of job stack z_0) with jobs $2, \ldots, i$ in fixed positions $m_2, \ldots, m_i$ of job stack z_0. Observe that the times at which V^0 hits the set U^i are a subsequence of the times at which V^0 hits the set U^{i-1}. Comparison of the marked job and labelled jobs methods rests on c.l.t.'s in continuous time for U^i blocks (defined by the successive exits from the set U^i) of V^0.

Denote by δ_k^i the length of the kth U^i block of V^0, $k \ge 1$. Let α_k^{ji} be the number of passage times for job j in the kth U^i block and let Y_k^{ji} be the sum of the values of the function f for these passage times, $1 \le j \le i$. Then set $e^{ji} = (E\{\delta_1^i\})^{1/2} \sigma^{ji} / E\{\alpha_1^{ji}\}$, where $(\sigma^{ji})^2 = \sigma^2(Y_k^{ji} - r(f)\alpha_k^{ji})$ for all k. We denote the analogous quantities for jobs $1, 2, \ldots, j$ by $\alpha_k^{(j)i}$, $Y_k^{(j)i}$, and $e^{(j)i} = (E\{\delta_1^{(j)i}\})^{1/2} \sigma^{(j)i} / E\{\alpha_1^{(j)i}\}$, respectively; here $(\sigma^{(j)i})^2 = \sigma^2(Y_k^{(j)i} - r(f)\alpha_k^{(j)i})$.

We require three lemmas. The first asserts that the statistical efficiency of a simulation based on observation of passage times for job 1 (the marked job) in U^1 blocks is the same as for a simulation based on observation of passage times for job i in U^i blocks, $1 \le i \le N$.

(4.8) LEMMA. $e^{ii} = e^{11}$.

Proof: Recall that $N^i(t)$ is the position of the job labelled i in the job stack at time t. Consider the process $X^i = \{X^i(t): t \ge 0\}$, where $X^i(t) = (Z(t), N^i(t))$. Let $L^i(t)$ be the last state occupied by X^i before jumping to $X^i(t)$ and set $V^i = \{V^i(t): t \ge 0\}$, where

$V^i(t) = (L^i(t), X^i(t))$. Let $\{P_n^i : n \geq 0\}$ be the sequence of passage times for job i enumerated in termination order. Observe that the successive entrances of V^i to $v_0 = (z_0', m_1', z_0, m_1)$ are regeneration points. Based on these v_0 cycles (for some e^i) by Lemma (4.1)

$$t^{1/2}\left(\frac{1}{m^i(t)}\sum_{n=1}^{m^i(t)} f(P_n^i) - r(f)\right)\Big/ e^i \Rightarrow N(0,1)$$

as $t \to \infty$, where $m^i(t)$ is the number of passage times for job i completed in $(0,t]$. Similarly, based on U^i blocks of V^i,

$$t^{1/2}\left(\frac{1}{m^i(t)}\sum_{n=1}^{m^i(t)} f(P_n^i) - r(f)\right)\Big/ e^{ii} \Rightarrow N(0,1) \tag{4.9}$$

as $t \to \infty$. Since the numerators and the limits are the same, the denominators must be the same; i.e., $e^{ii} = e^i$.

The desired result follows since $e^i = e^{11}$. To see this set $\zeta_1^i = \inf\{t > 0 : V^i(t) = v_0\}$ and consider the process $W^i = \{W^i(t) : t \geq 0\}$, where $W^i(t) = V^i(\zeta_1^i + t)$. Both V and W^i are generalized semi-Markov processes. These two generalized semi-Markov processes have the same probability mass functions and clock setting distributions since these building blocks are functions only of the routing matrix and the parameters of the service time distributions, and the jobs are stochastically identical. This implies that V and W^i have the same finite dimensional distributions. Enumerate the passage times for job i from time ζ_1^i as $\{R_n^i : n \geq 0\}$ and let $m^{i\prime}(t)$ be the number of passage times for job i completed in $(\zeta_1^i, \zeta_1^i + t]$. Denote the sequence of passage times for job 1 by $\{P_n^1 : n \geq 0\}$ and let $m^1(t)$ be the number of passage times for job 1 completed in $(0,t]$. Now consider the functional $\{C^1(t) : t \geq 0\}$ of V where

$$C^1(t) = \frac{1}{m^1(t)} \sum_{n=1}^{m^1(t)} f(P_n^1) - r(f)$$

and observe that $\{C^i(t):t\geq 0\}$, where

$$C^i(t) = \frac{1}{m^i(t)} \sum_{n=1}^{m^i(t)} f(R_n^i) - r(f)$$

is the same functional of the process W^i. Thus for all $t>0$, $C^1(t)$ and $C^i(t)$ have the same distribution. Since

$$t^{1/2}\ C^1(t)/e^{11} \Rightarrow N(0,1)$$

as $t\to\infty$, and

$$t^{1/2}\ C^i(t)/e^i \Rightarrow N(0,1)$$

as $t\to\infty$ by Lemma (4.1), $e^i = e^{11}$. □

Next we show that simulations based on observation of passage times for jobs $1,2,\ldots,j$ in U^i blocks and on passage times for jobs $1,2,\ldots,j$ in U^{i+1} blocks are equally efficient.

(4.10) LEMMA. Let $l = i + 1$. Then $e^{(j)i} = e^{(j)l}$ for $j=1,2,\ldots,i$.

Proof: Set $v_0^{(k)} = (z_0',m_1',m_2',\ldots,m_k',z_0,m_1,m_2,\ldots,m_k)$, $1\leq k\leq N$. Consider the process $X^{(k)} = \{X^{(k)}(t):t\geq 0\}$, where $X^{(k)}(t) = (Z(t),N^1(t),\ldots,N^k(t))$. Let $L^{(k)}(t)$ be the last state occupied by $X^{(k)}$ before jumping to $X^{(k)}(t)$. Set $V^{(k)} = \{V^{(k)}(t):t\geq 0\}$, where $V^{(k)}(t) = (L^{(k)}(t),X^{(k)}(t))$. The successive entrances to state $v_0^{(i)}$ are regeneration points for the process $V^{(i)}$ and the successive entrances to state $v_0^{(l)}$ are regeneration points for $V^{(l)}$. Let $\{P_n^{(j)}:n\geq 0\}$ be the sequence of passage times for jobs $1,2,\ldots,j$ enumerated in termination order. Then

$$t^{1/2}\left(\frac{1}{m^{(j)}(t)}\sum_{n=1}^{m^{(j)}(t)} f(P_n^{(j)}) - r(f)\right)\Big/ e^{(j)t} \Rightarrow N(0,1)$$

as $t \to \infty$ and

$$t^{1/2}\left(\frac{1}{m^{(j)}(t)}\sum_{n=1}^{m^{(j)}(t)} f(P_n^{(j)}) - r(f)\right)\Big/ e^{(j)t} \Rightarrow N(0,1)$$

as $t \to \infty$, where $m^{(j)}(t)$ is the number of passage times for jobs $1,2,\ldots,j$ completed in $(0,t]$. Since the numerators and the limits are the same, the denominators must be the same. □

Lemma (4.11) asserts that a simulation based on observation of passage times for jobs 1 and 2 in U^2 blocks is at least as efficient as a simulation based on passage times for job 1 in U^1 blocks. The idea is to develop a bivariate c.l.t. for passage times of jobs 1 and 2 in U^2 blocks, apply the continuous mapping theorem to obtain a one dimensional c.l.t., and make an identification of the resulting variance constant.

(4.11) LEMMA. $e^{(2)2} \le e^{11}$.

Proof: Let W_k be the column vector with components W_k^1 and W_k^2, where $W_k^1 = Y_k^{12} - r(f)\alpha_k^{12}$ and $W_k^2 = Y_k^{22} - r(f)\alpha_k^{22}$. Then the vectors $\{W_k : k \ge 0\}$ are i.i.d. with mean 0. It follows that

$$n^{-1/2} \sum_{k=1}^{n} W_k \Rightarrow N(0,\Sigma),$$

as $n \to \infty$, where $\Sigma = \{\sigma_{ij}\}$ and $\sigma_{ij} = E\{W_1^i W_1^j\}$ for $i,j = 1,2$. An application of the continuous mapping theorem using the mapping $h(w_1,w_2) = w_1 + w_2$ yields

$$n^{-1/2} \sum_{k=1}^{n} \sum_{i=1}^{2} W_k^i \Rightarrow N(0,\sigma^2),$$

as $n \to \infty$, where

$$\sigma^2 = \sum_i \sigma_{ii} + 2 \sum_{i<j} \sigma_{ij}.$$

Furthermore from the form of this c.l.t. (since $W_k^1 + W_k^2$ corresponds to jobs 1 and 2), it is clear that $\sigma^2 = (\sigma^{(2)2})^2$, $\sigma_{11} = (\sigma^{12})^2$, and $\sigma_{22} = (\sigma^{22})^2$.

Next observe that $e^{22} = e^{11}$ by Lemma (4.8) and $e^{12} = e^{11}$ by Lemma (4.10). Using the definitions of e^{12}, e^{11}, and e^{22}, these results imply

$$\sigma_{11} = (\sigma^{11})^2 E\{\delta_1^1\}\left(E\{\alpha_1^{12}\}\right)^2 \Big/ \left[E\{\delta_1^2\}\left(E\{\alpha_1^{11}\}\right)^2\right]$$

and

$$\sigma_{22} = (\sigma^{11})^2 E\{\delta_1^1\}\left(E\{\alpha_1^{22}\}\right)^2 \Big/ \left[E\{\delta_1^2\}\left(E\{\alpha_1^{11}\}\right)^2\right].$$

From the Cauchy-Schwarz inequality we have

$$2|\sigma_{12}| \le 2(\sigma_{11}\sigma_{22})^{1/2} = (\sigma^{11})^2 E\{\delta_1^1\}\left(2E\{\alpha_1^{22}\}E\{\alpha_1^{12}\}\right) \Big/ \left[E\{\delta_1^2\}\left(E\{\alpha_1^{11}\}\right)^2\right].$$

Hence

$$\begin{aligned}(\sigma^{(2)2})^2 &= \sigma_{11} + \sigma_{22} + 2\sigma_{12} \\ &\le \left[(\sigma^{11})^2 E\{\delta_1^1\}\left(E\{\alpha_1^{22}\} + E\{\alpha_1^{12}\}\right)^2\right] \Big/ \left[E\{\delta_1^2\}\left(E\{\alpha_1^{11}\}\right)^2\right].\end{aligned}$$

Since $E\{\alpha_1^{(2)2}\} = E\{\alpha_1^{22}\} + E\{\alpha_1^{12}\}$, $e^{(2)2}/e^{11} \le 1$. □

(4.12) PROPOSITION. For all functions f, $e^0(f) \le e(f)$.

Proof: It is sufficient to show that $e^{(N),N} \le e^{11}$. To see this first observe (using Lemma (4.1) again) that

$$t^{1/2}\left(\frac{1}{m^0(t)}\sum_{n=1}^{m^0(t)} f(P_n^0) - r(f)\right)\Big/ e^{(N)N} \Rightarrow N(0,1)$$

as $t \to \infty$ and

$$t^{1/2}\left(\frac{1}{m^0(t)}\sum_{n=1}^{m^0(t)} f(P_n^0) - r(f)\right)\Big/ e^0(f) \Rightarrow N(0,1)$$

as $t \to \infty$, where $m^0(t)$ is the number of passage times (irrespective of job identity) completed in $(0,t]$; therefore $e^{(N)N} = e^0(f)$. But $e(f) = e^{11}$; this follows from the c.l.t.'s of Equations (4.6) and (4.9) (with $i=1$).

We first show that $e^{(3)3} \le e^{11}$. Let W_k be the column vector with components W_k^1 and W_k^2, where $W_k^1 = Y_k^{(2)3} - r(f)\alpha_k^{(2)3}$ and $W_k^2 = Y_k^{33} - r(f)\alpha_k^{33}$. Then the vectors $\{W_k : k \ge 0\}$ are i.i.d. with mean 0 and therefore

$$n^{-1/2}\sum_{k=1}^{n} W_k \Rightarrow N(0,\Sigma),$$

as $n \to \infty$, where $\Sigma = \{\sigma_{ij}\}$ and $\sigma_{ij} = E\{W_1^i W_1^j\}$. By the argument used in the proof of Lemma (4.11), $\sigma_{11} = (\sigma^{(2)3})^2$, and $\sigma_{22} = (\sigma^{33})^2$. By Lemma (4.8), $e^{33} = e^{11}$. Therefore

$$\sigma_{22} = (\sigma^{11})^2 E\{\delta_1^1\}\left(E\{\alpha_1^{33}\}\right)^2 \Big/ \left[E\{\delta_1^3\}\left(E\{\alpha_1^{11}\}\right)^2\right].$$

Next observe that $e^{(2)2} \le e^{11}$ by Lemma (4.11) and $e^{(2)3} = e^{(2)2}$ by Lemma (4.10). Thus

$$\sigma_{11} \le (\sigma^{11})^2 E\{\delta_1^1\}\left(E\{\alpha_1^{(2)3}\}\right)^2 \Big/ \left[E\{\delta_1^3\}\left(E\{\alpha_1^{11}\}\right)^2\right].$$

By the Cauchy-Schwarz inequality,

$$(\sigma^{(3)2})^2 \le \sigma_{11} + \sigma_{22}^2 + 2(\sigma_{11}\sigma_{22})^{1/2}.$$

Using the fact that $E\{\alpha_1^{(3)3}\} = E\{\alpha_1^{33}\} + E\{\alpha_1^{(2)3}\}$, it follows that $e^{(3)3}/e^{11} \leq 1$. The same argument (using $e^{(3)3} \leq e^{11}$) shows that $e^{(4)4} \leq e^{11}$. Continuing in this way, it follows that $e^{(N)N} \leq e^{11}$. □

5. CONCLUDING REMARKS

The assumption that service times at any center have a density function that is positive and continuous on $(0,\infty)$ guarantees (for networks with single states) that $\{(X(S_n),P_{n+1}):n \geq 0\}$ and $\{(Z(T_n^0),P_{n+1}^0):n \geq 0\}$ are regenerative processes in discrete time. The proof that $e^0(f) \leq e(f)$ requires only that the proceses be regenerative.

ACKNOWLEDGMENT

We are both grateful to the National Science Foundation for support under Grant MCS-7909139. In addition, Donald L. Iglehart gratefully acknowledges partial support under Office of Naval Research Contract N00014-76-C-0578 (NR 042-343). We thank the referee for the suggestions which led to an improved exposition.

REFERENCES

[1] Billingsley, P. (1968). *Convergence of Probability Measures.* John Wiley. New York.

[2] Billingsley, P. (1979). *Probability and Measure.* John Wiley. New York.

[3] Chung, K. L. (1974). *A Course in Probability Theory.* Second Edition. Academic Press. New York.

[4] Crane, M. A. and Iglehart, D. L. (1975). Simulating stable stochastic systems: III, Regenerative processes and discrete event simulation. *Operations Research 23,* 33-45.

[5] Hordijk, A., Iglehart, D. L. and Schassberger, R. (1976). Discrete time methods for simulating continuous time Markov chains. *Advances in Appl. Probability 8,* 772-778.

[6] Iglehart, D. L. (1971). Functional limit theorems for the queue GI/G/1 in light traffic. *Advances in Appl. Probability 3,* 269-281.

[7] Iglehart, D. L. and Shedler, G. S. (1980). *Regenerative simulation of response times in networks of queues.* Lecture Notes in Control and Information Sciences, vol. 26. Springer-Verlag. Berlin, Heidelberg, New York.

[8] Iglehart, D. L. and Shedler, G. S. (1981). Regenerative simulation of response times in networks of queues: statistical efficiency. *Acta Informatica 15,* 347-363.

[9] Iglehart, D. L. and Shedler, G. S. (1981). Simulation for passage times in closed, multiclass networks of queues with general service times. IBM Research Report RJ3191. San Jose, California.

[10] Shedler, G.S. and Slutz, D. R. (1981). Irreducibility in closed multiclass networks of queues with priorities: passage times of a marked job. *Performance Evaluation 1,* 334-343.

[11] Shedler, G. S. and Southard, J. (1981). Simulation for passage times in closed, multiclass networks of queues with unrestricted priorities. IBM Research Report RJ3190. San Jose, California.

[12] Shedler, G. S. and Southard, J. (1981). Regenerative simulation of networks of queues with general service times: passage through subnetworks. IBM Research Report RJ3221. San Jose, California. To appear in *IBM J. Res. Develop.*

[13] Whitt, W. (1980). Continuity of generalized semi-Markov processes. *Math. Oper. Res. 5,* 494-501.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 58	2. GOVT ACCESSION NO. AD-A119154	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) STATISTICAL EFFICIENCY OF REGENERATIVE SIMULATION METHODS FOR NETWORKS OF QUEUES		5. TYPE OF REPORT & PERIOD COVERED TECHNICAL REPORT 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) DONALD L. IGLEHART GERALD S. SHEDLER		8. CONTRACT OR GRANT NUMBER(s) N00014-76-C-0678
9. PERFORMING ORGANIZATION NAME AND ADDRESS DEPARTMENT OF OPERATIONS RESEARCH STANFORD UNIVERSITY STANFORD, CA 94305		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS (NR 042-343)
11. CONTROLLING OFFICE NAME AND ADDRESS STATISTICS AND PROBABILITY PROGRAM OFFICE OF NAVAL RESEARCH (CODE 436) ARLINGTON, VA 20360		12. REPORT DATE MARCH 1982 13. NUMBER OF PAGES 23
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

THIS DOCUMENT HAS BEEN APPROVED FOR PUBLIC RELEASE AND SALE; ITS DISTRIBUTION IS UNLIMITED.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

generalized semi-Markov processes
networks of queues
passage times
regenerative simulation
simulation output analysis

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

SEE ATTACHED

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

ABSTRACT: This paper is concerned with the assessment of the statistical efficiency of proposed regenerative simulation methods. We compare the efficiency of the "marked job" and "labelled jobs" methods for estimation of passage times in multiclass networks of queues with general service times. Using central limit theorem arguments, we show that the confidence intervals constructed for the expected value of a general function of the limiting passage time using the labelled jobs method are shorter than those obtained from the marked job method. This is consistent with intuition since the labelled jobs method extracts more passage time information from a fixed length simulation run.

FILME

-8